



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/944,919	08/31/2001	Maria Castellanos	HP-10007912	3894

7590 09/28/2004

HEWLETT-PACKARD COMPANY  
Intellectual Property Administration  
P.O. Box 272400  
Fort Collins, CO 80527-2400

EXAMINER
----------

AMSBURY, WAYNE P

ART UNIT	PAPER NUMBER
----------	--------------

2171

DATE MAILED: 09/28/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

**Office Action Summary**

Application No.

09/944,919

Applicant(s)

CASTELLANOS ET AL.

Examiner

Wayne Amsbury

Art Unit

2171

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

- 1) ☒ Responsive to communication(s) filed on 22 July 2004.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

- 4) ☒ Claim(s) 1-12 and 15-35 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) 31-35 is/are allowed.
- 6) ☒ Claim(s) 1-6, 9-12, 15-16, 19-26 and 29-30 is/are rejected.
- 7) ☐ Claim(s) 7,8,17,18,27 and 28 is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

**Application Papers**

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 31 August 2001 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some \* c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
  2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

- |  |   |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)  | 4) <input type="checkbox"/> Interview Summary (PTO-413)<br>Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)                                   | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152)             |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)<br>Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____  |

CLAIMS 1-12 AND 15-35 ARE PENDING

1. The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

2. The finality of the previous action is hereby withdrawn in light of prior art discovered with an update search.

3. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

**Claims 1-2, 5 and 9-10 are rejected under 35 U.S.C. 102(e) as being anticipated by Thompson et al (Thompson), US 2002/0103834, 1 August 2002.**

Thompson is directed to analyzing documents in electronic form in several stages that include cleaning text images, error correction (cleaning) of ASCII text, and data mining of the cleaned text [FIG 1].

It is considered that this corresponds to a computer-implemented method for mining a document containing dirty text in order to profile (summarize) a document [0860]. It is further considered that some of the claims are clearly anticipated by Thompson, but in the interest of compact prosecution the details of the teachings of Thompson are set forth below.

As to **claim 1**, in re FIG 1, stage 3 of the image-to-data processing produces cleaned text images from text images corresponds to removing an instance of dirty text within a document to produce a cleaned document having a content, and so does stage 2 of the text-to-data conversion, where the error correction corresponds to cleaning dirty text. (See Image evaluation beginning at [0425] for further details.) Stage 3 of the text-to-data conversion performs a data mining operation on a cleaned document, and the content analysis derives further relevant information from the cleaned document and provides a profile that corresponds to a summary of the content of the document. See [0860], [0882] and [0895] for details, and in particular, note that *abstracting* is used as a means of summarizing a document's content.

As to **claim 2**, Thompson corrects errors in text including misspelling and grammatical errors according to instructions provided by the user [0476; 0480; 0484].

As to **claim 5**, Thompson treats sentences as a form of grammatical set, which is identified by identifying a beginning and an end [0921-0922].

Art Unit: 2171

As to **claims 9-10**, the user provides text correction according to instructions provided by the user [0480], including the selection of working documents [0482], and dictionaries [0484].

This corresponds to selecting a text-mining component. This involves at least the assignment of a parameter representing a confidence value.

**4. Claims 3-4, 6, 11-12, 15-16, 19-20, 21-26 and 29-30 are rejected under 35 U.S.C. 103(a) as being unpatentable over Thompson et al (Thompson), US 2002/0103834, 1 August 2002.**

As to **claim 3**, Thompson includes documents in the computer category in the subject categories of interest [0058]. **Official Notice** is taken that documents in the computer subject category contain computer code. **It would have been obvious** to one of ordinary skill in the art at the time of the invention to remove an instance of computer code from a document because this is a structural element that may not be but may need to be entirely correct.

As to **claim 4**, Thompson applies swap-out tables (substitution tables) of arbitrary size [0734]-[0740]. Clearly a table of a single entry or a table, all of whose entries are swapped, corresponds to removing a table from a document and replacing it with another. It should be noted that Thompson also determines that documents do or do not contain tables of their own [0970].

Art Unit: 2171

Thompson also screens and may remove documents that are either legal forms [0891] or medical forms [0889], either of which are well known to contain or be tables. Thompson mines data from non-standardized formats [0029] that include forms. To the extent that Thompson does not anticipate removing a table from a document:

**Official Notice** is taken that tables were well known at the time of the invention as common structural elements in documents and as forms. **It would have been obvious** to one of ordinary skill in the art at the time of the invention to include tables as document components to be corrected and analyzed because they are recognized [0970] for no additional cost, they may contain errors like other structural components, and they can be corrected by the techniques of Thompson.

As to **claim 6**, Thompson assigns a quality rating to each working document, and such a rating corresponds to both a scoring and a ranking [0490]-[0500]. However, he does not explicitly apply the rating system to components such as grammatical sets and/or sentences, even though these components are analyzed prior to addition to the vocabulary [00928]. **It would have been obvious** to one of ordinary skill in the art at the time of the invention to apply the quality ranking of Thompson to grammatical sets because this would allow the user to decide whether or not to index them [0931], thereby reducing the number of reference databases.

The elements of **claims 11-12, 15-16, 19-20, 21-26 and 28-29** are rejected in the analysis above and these claims are rejected on that basis.

Art Unit: 2171

**5. Claims 7-8, 17-18 and 27-28** are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

The particular scoring and summarizing technique set forth in these claims in the context of the other limitations is neither anticipated nor suggested by the prior art of record.

**6. Claims 31-35 are allowed.**

The combination of elements of these claims, including scoring and ranking sentences in combination with removing tables and computer code is neither anticipated nor suggested by the prior art of record.

**7.** Any inquiry concerning this communication or earlier communications from the examiner should be directed to Wayne Amsbury whose telephone number is 703-305-3828. The examiner can normally be reached on M-TH 7-5.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Safet Metjahic can be reached on 703-308-1436. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Art Unit: 2171

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

WPA

  
WAYNE AMSBURY  
PRIMARY PATENT EXAMINER